

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/558, 149
Source: TFW16
Date Processed by STIC: 10/03/2005

ENTERED



IFW16

RAW SEQUENCE LISTING DATE: 10/03/2005
 PATENT APPLICATION: US/09/558,149 TIME: 10:58:37

Input Set : N:\Cr3f3\RULE60\09558149.raw
 Output Set: N:\CRF4\09292005\I558149.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:
 5 (i) APPLICANT: NICOLAIDES, NICHOLAS
 6 VOGELSTEIN, BERT
 7 KINZLER, KINZLER
 9 (ii) TITLE OF INVENTION: A METHOD FOR GENERATING HYPERMUTABLE
 10 ORGANISMS
 12 (iii) NUMBER OF SEQUENCES: 2
 14 (iv) CORRESPONDENCE ADDRESS:
 15 (A) ADDRESSEE: Banner & Witcoff
 16 (B) STREET: 1001 G Street, NW
 17 (C) CITY: Washington
 18 (D) STATE: DC
 19 (E) COUNTRY: USA
 20 (F) ZIP: 20001
 22 (v) COMPUTER READABLE FORM:
 23 (A) MEDIUM TYPE: Diskette
 24 (B) COMPUTER: IBM Compatible
 25 (C) OPERATING SYSTEM: DOS
 26 (D) SOFTWARE: FastSEQ for Windows Version 2.0
 28 (vi) CURRENT APPLICATION DATA:
 C--> 29 (A) APPLICATION NUMBER: US/09/558,149
 C--> 30 (B) FILING DATE: 26-Apr-2000
 W--> 36 (C) CLASSIFICATION: 800
 33 (vii) PRIOR APPLICATION DATA:
 34 (A) APPLICATION NUMBER: US/09/059,461
 35 (B) FILING DATE: 14-APRIL-1998
 39 (viii) ATTORNEY/AGENT INFORMATION:
 40 (A) NAME: Kagan, Sarah A
 41 (B) REGISTRATION NUMBER: 32141
 42 (C) REFERENCE/DOCKET NUMBER: 01107.73306
 44 (ix) TELECOMMUNICATION INFORMATION:
 45 (A) TELEPHONE: 202-508-9100
 46 (B) TELEFAX: 202-508-9299
 47 (C) TELEX:
 50 (2) INFORMATION FOR SEQ ID NO: 1:
 52 (i) SEQUENCE CHARACTERISTICS:
 53 (A) LENGTH: 2771 base pairs
 54 (B) TYPE: nucleic acid
 55 (C) STRANDEDNESS: double
 56 (D) TOPOLOGY: linear
 58 (ii) MOLECULE TYPE: cDNA
 59 (ix) FEATURE:

RAW SEQUENCE LISTING

DATE: 10/03/2005

PATENT APPLICATION: US/09/558,149

TIME: 10:58:37

Input Set : N:\Crf3\RULE60\09558149.raw

Output Set: N:\CRF4\09292005\I558149.raw

```

61      (A) NAME/KEY: Coding Sequence
62      (B) LOCATION: 25...2610
63      (D) OTHER INFORMATION:
66      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
68      CGAGGCGGAT CGGGTGTTC ATCC ATG GAG CGA GCT GAG AGC TCG AGT ACA      51
69      Met Glu Arg Ala Glu Ser Ser Ser Thr
70      1 5
72      GAA CCT GCT AAG GCC ATC AAA CCT ATT GAT CGG AAG TCA GTC CAT CAG      99
73      Glu Pro Ala Lys Ala Ile Lys Pro Ile Asp Arg Lys Ser Val His Gln
74      10 15 20 25
76      ATT TGC TCT GGG CAG GTG GTA CTG AGT CTA AGC ACT GCG GTA AAG GAG      147
77      Ile Cys Ser Gly Gln Val Val Leu Ser Leu Ser Thr Ala Val Lys Glu
78      30 35 40
80      TTA GTA GAA AAC AGT CTG GAT GCT GGT GCC ACT AAT ATT GAT CTA AAG      195
81      Leu Val Glu Asn Ser Leu Asp Ala Gly Ala Thr Asn Ile Asp Leu Lys
82      45 50 55
84      CTT AAG GAC TAT GGA GTG GAT CTT ATT GAA GTT TCA GAC AAT GGA TGT      243
85      Leu Lys Asp Tyr Gly Val Asp Leu Ile Glu Val Ser Asp Asn Gly Cys
86      60 65 70
88      GGG GTA GAA GAA GAA AAC TTC GAA GGC TTA ACT CTG AAA CAT CAC ACA      291
89      Gly Val Glu Glu Glu Asn Phe Glu Gly Leu Thr Leu Lys His His Thr
90      75 80 85
92      TCT AAG ATT CAA GAG TTT GCC GAC CTA ACT CAG GTT GAA ACT TTT GGC      339
93      Ser Lys Ile Gln Glu Phe Ala Asp Leu Thr Gln Val Glu Thr Phe Gly
94      90 95 100 105
96      TTT CGG GGG GAA GCT CTG AGC TCA CTT TGT GCA CTG AGC GAT GTC ACC      387
97      Phe Arg Gly Glu Ala Leu Ser Ser Leu Cys Ala Leu Ser Asp Val Thr
98      110 115 120
100     ATT TCT ACC TGC CAC GCA TCG GCG AAG GTT GGA ACT CGA CTG ATG TTT      435
101     Ile Ser Thr Cys His Ala Ser Ala Lys Val Gly Thr Arg Leu Met Phe
102     125 130 135
104     GAT CAC AAT GGG AAA ATT ATC CAG AAA ACC CCC TAC CCC CGC CCC AGA      483
105     Asp His Asn Gly Lys Ile Ile Gln Lys Thr Pro Tyr Pro Arg Pro Arg
106     140 145 150
108     GGG ACC ACA GTC AGC GTG CAG CAG TTA TTT TCC ACA CTA CCT GTG CGC      531
109     Gly Thr Thr Val Ser Val Gln Gln Leu Phe Ser Thr Leu Pro Val Arg
110     155 160 165
112     CAT AAG GAA TTT CAA AGG AAT ATT AAG AAG GAG TAT GCC AAA ATG GTC      579
113     His Lys Glu Phe Gln Arg Asn Ile Lys Lys Glu Tyr Ala Lys Met Val
114     170 175 180 185
116     CAG GTC TTA CAT GCA TAC TGT ATC ATT TCA GCA GGC ATC CGT GTA AGT      627
117     Gln Val Leu His Ala Tyr Cys Ile Ile Ser Ala Gly Ile Arg Val Ser
118     190 195 200
120     TGC ACC AAT CAG CTT GGA CAA GGA AAA CGA CAG CCT GTG GTA TGC ACA      675
121     Cys Thr Asn Gln Leu Gly Gln Gly Lys Arg Gln Pro Val Val Cys Thr
122     205 210 215
124     GGT GGA AGC CCC AGC ATA AAG GAA AAT ATC GGC TCT GTG TTT GGG CAG      723
125     Gly Gly Ser Pro Ser Ile Lys Glu Asn Ile Gly Ser Val Phe Gly Gln
126     220 225 230

```

RAW SEQUENCE LISTING

DATE: 10/03/2005

PATENT APPLICATION: US/09/558,149

TIME: 10:58:37

Input Set : N:\Crif3\RULE60\09558149.raw

Output Set: N:\CRF4\09292005\I558149.raw

128	AAG CAG TTG CAA AGC CTC ATT CCT TTT GTT CAG CTG CCC CCT AGT GAC	771
129	Lys Gln Leu Gln Ser Leu Ile Pro Phe Val Gln Leu Pro Pro Ser Asp	
130	235 240 245	
132	TCC GTG TGT GAA GAG TAC GGT TTG AGC TGT TCG GAT GCT CTG CAT AAT	819
133	Ser Val Cys Glu Glu Tyr Gly Leu Ser Cys Ser Asp Ala Leu His Asn	
134	250 255 260 265	
136	CTT TTT TAC ATC TCA GGT TTC ATT TCA CAA TGC ACG CAT GGA GTT GGA	867
137	Leu Phe Tyr Ile Ser Gly Phe Ile Ser Gln Cys Thr His Gly Val Gly	
138	270 275 280	
140	AGG AGT TCA ACA GAC AGA CAG TTT TTC TTT ATC AAC CGG CGG CCT TGT	915
141	Arg Ser Ser Thr Asp Arg Gln Phe Phe Phe Ile Asn Arg Arg Pro Cys	
142	285 290 295	
144	GAC CCA GCA AAG GTC TGC AGA CTC GTG AAT GAG GTC TAC CAC ATG TAT	963
145	Asp Pro Ala Lys Val Cys Arg Leu Val Asn Glu Val Tyr His Met Tyr	
146	300 305 310	
148	AAT CGA CAC CAG TAT CCA TTT GTT GTT CTT AAC ATT TCT GTT GAT TCA	1011
149	Asn Arg His Gln Tyr Pro Phe Val Val Leu Asn Ile Ser Val Asp Ser	
150	315 320 325	
152	GAA TGC GTT GAT ATC AAT GTT ACT CCA GAT AAA AGG CAA ATT TTG CTA	1059
153	Glu Cys Val Asp Ile Asn Val Thr Pro Asp Lys Arg Gln Ile Leu Leu	
154	330 335 340 345	
156	CAA GAG GAA AAG CTT TTG TTG GCA GTT TTA AAG ACC TCT TTG ATA GGA	1107
157	Gln Glu Glu Lys Leu Leu Leu Ala Val Leu Lys Thr Ser Leu Ile Gly	
158	350 355 360	
160	ATG TTT GAT AGT GAT GTC AAC AAG CTA AAT GTC AGT CAG CAG CCA CTG	1155
161	Met Phe Asp Ser Asp Val Asn Lys Leu Asn Val Ser Gln Gln Pro Leu	
162	365 370 375	
164	CTG GAT GTT GAA GGT AAC TTA ATA AAA ATG CAT GCA GCG GAT TTG GAA	1203
165	Leu Asp Val Glu Gly Asn Leu Ile Lys Met His Ala Ala Asp Leu Glu	
166	380 385 390	
168	AAG CCC ATG GTA GAA AAG CAG GAT CAA TCC CCT TCA TTA AGG ACT GGA	1251
169	Lys Pro Met Val Glu Lys Gln Asp Gln Ser Pro Ser Leu Arg Thr Gly	
170	395 400 405	
172	GAA GAA AAA AAA GAC GTG TCC ATT TCC AGA CTG CGA GAG GCC TTT TCT	1299
173	Glu Glu Lys Lys Asp Val Ser Ile Ser Arg Leu Arg Glu Ala Phe Ser	
174	410 415 420 425	
176	CTT CGT CAC ACA ACA GAG AAC AAG CCT CAC AGC CCA AAG ACT CCA GAA	1347
177	Leu Arg His Thr Thr Glu Asn Lys Pro His Ser Pro Lys Thr Pro Glu	
178	430 435 440	
180	CCA AGA AGG AGC CCT CTA GGA CAG AAA AGG GGT ATG CTG TCT TCT AGC	1395
181	Pro Arg Arg Ser Pro Leu Gly Gln Lys Arg Gly Met Leu Ser Ser Ser	
182	445 450 455	
184	ACT TCA GGT GCC ATC TCT GAC AAA GGC GTC CTG AGA CCT CAG AAA GAG	1443
185	Thr Ser Gly Ala Ile Ser Asp Lys Gly Val Leu Arg Pro Gln Lys Glu	
186	460 465 470	
188	GCA GTG AGT TCC AGT CAC GGA CCC AGT GAC CCT ACG GAC AGA GCG GAG	1491
189	Ala Val Ser Ser Ser His Gly Pro Ser Asp Pro Thr Asp Arg Ala Glu	
190	475 480 485	
192	GTG GAG AAG GAC TCG GGG CAC GGC AGC ACT TCC GTG GAT TCT GAG GGG	1539

RAW SEQUENCE LISTING

DATE: 10/03/2005

PATENT APPLICATION: US/09/558,149

TIME: 10:58:37

Input Set : N:\Crf3\RULE60\09558149.raw

Output Set: N:\CRF4\09292005\I558149.raw

193	Val	Glu	Lys	Asp	Ser	Gly	His	Gly	Ser	Thr	Ser	Val	Asp	Ser	Glu	Gly	
194	490					495					500					505	
196	TTC	AGC	ATC	CCA	GAC	ACG	GGC	AGT	CAC	TGC	AGC	AGC	GAG	TAT	GCG	GCC	1587
197	Phe	Ser	Ile	Pro	Asp	Thr	Gly	Ser	His	Cys	Ser	Ser	Glu	Tyr	Ala	Ala	
198					510					515					520		
200	AGC	TCC	CCA	GGG	GAC	AGG	GGC	TCG	CAG	GAA	CAT	GTG	GAC	TCT	CAG	GAG	1635
201	Ser	Ser	Pro	Gly	Asp	Arg	Gly	Ser	Gln	Glu	His	Val	Asp	Ser	Gln	Glu	
202				525					530					535			
204	AAA	GCG	CCT	GAA	ACT	GAC	GAC	TCT	TTT	TCA	GAT	GTG	GAC	TGC	CAT	TCA	1683
205	Lys	Ala	Pro	Glu	Thr	Asp	Asp	Ser	Phe	Ser	Asp	Val	Asp	Cys	His	Ser	
206			540					545				550					
208	AAC	CAG	GAA	GAT	ACC	GGA	TGT	AAA	TTT	CGA	GTT	TTG	CCT	CAG	CCA	ACT	1731
209	Asn	Gln	Glu	Asp	Thr	Gly	Cys	Lys	Phe	Arg	Val	Leu	Pro	Gln	Pro	Thr	
210		555				560				565							
212	AAT	CTC	GCA	ACC	CCA	AAC	ACA	AAG	CGT	TTT	AAA	AAA	GAA	GAA	ATT	CTT	1779
213	Asn	Leu	Ala	Thr	Pro	Asn	Thr	Lys	Arg	Phe	Lys	Lys	Glu	Glu	Ile	Leu	
214	570				575					580					585		
216	TCC	AGT	TCT	GAC	ATT	TGT	CAA	AAG	TTA	GTA	AAT	ACT	CAG	GAC	ATG	TCA	1827
217	Ser	Ser	Ser	Asp	Ile	Cys	Gln	Lys	Leu	Val	Asn	Thr	Gln	Asp	Met	Ser	
218			590						595					600			
220	GCC	TCT	CAG	GTT	GAT	GTA	GCT	GTG	AAA	ATT	AAT	AAG	AAA	GTT	GTG	CCC	1875
221	Ala	Ser	Gln	Val	Asp	Val	Ala	Val	Lys	Ile	Asn	Lys	Lys	Val	Val	Pro	
222			605					610					615				
224	CTG	GAC	TTT	TCT	ATG	AGT	TCT	TTA	GCT	AAA	CGA	ATA	AAG	CAG	TTA	CAT	1923
225	Leu	Asp	Phe	Ser	Met	Ser	Ser	Leu	Ala	Lys	Arg	Ile	Lys	Gln	Leu	His	
226			620					625				630					
228	CAT	GAA	GCA	CAG	CAA	AGT	GAA	GGG	GAA	CAG	AAT	TAC	AGG	AAG	TTT	AGG	1971
229	His	Glu	Ala	Gln	Gln	Ser	Glu	Gly	Glu	Gln	Asn	Tyr	Arg	Lys	Phe	Arg	
230		635				640				645							
232	GCA	AAG	ATT	TGT	CCT	GGA	GAA	AAT	CAA	GCA	GCC	GAA	GAT	GAA	CTA	AGA	2019
233	Ala	Lys	Ile	Cys	Pro	Gly	Glu	Asn	Gln	Ala	Ala	Glu	Asp	Glu	Leu	Arg	
234	650				655				660						665		
236	AAA	GAG	ATA	AGT	AAA	ACG	ATG	TTT	GCA	GAA	ATG	GAA	ATC	ATT	GGT	CAG	2067
237	Lys	Glu	Ile	Ser	Lys	Thr	Met	Phe	Ala	Glu	Met	Glu	Ile	Ile	Gly	Gln	
238			670					675						680			
240	TTT	AAC	CTG	GGA	TTT	ATA	ATA	ACC	AAA	CTG	AAT	GAG	GAT	ATC	TTC	ATA	2115
241	Phe	Asn	Leu	Gly	Phe	Ile	Ile	Thr	Lys	Leu	Asn	Glu	Asp	Ile	Phe	Ile	
242			685					690					695				
244	GTG	GAC	CAG	CAT	GCC	ACG	GAC	GAG	AAG	TAT	AAC	TTC	GAG	ATG	CTG	CAG	2163
245	Val	Asp	Gln	His	Ala	Thr	Asp	Glu	Lys	Tyr	Asn	Phe	Glu	Met	Leu	Gln	
246			700					705				710					
248	CAG	CAC	ACC	GTG	CTC	CAG	GGG	CAG	AGG	CTC	ATA	GCA	CCT	CAG	ACT	CTC	2211
249	Gln	His	Thr	Val	Leu	Gln	Gly	Gln	Arg	Leu	Ile	Ala	Pro	Gln	Thr	Leu	
250		715				720				725							
252	AAC	TTA	ACT	GCT	GTT	AAT	GAA	GCT	GTT	CTG	ATA	GAA	AAT	CTG	GAA	ATA	2259
253	Asn	Leu	Thr	Ala	Val	Asn	Glu	Ala	Val	Leu	Ile	Glu	Asn	Leu	Glu	Ile	
254	730				735				740					745			
256	TTT	AGA	AAG	AAT	GGC	TTT	GAT	TTT	GTT	ATC	GAT	GAA	AAT	GCT	CCA	GTC	2307
257	Phe	Arg	Lys	Asn	Gly	Phe	Asp	Phe	Val	Ile	Asp	Glu	Asn	Ala	Pro	Val	

RAW SEQUENCE LISTING

DATE: 10/03/2005

PATENT APPLICATION: US/09/558,149

TIME: 10:58:37

Input Set : N:\Crif3\RULE60\09558149.raw

Output Set: N:\CRF4\09292005\I558149.raw

```

258          750          755          760
260 ACT GAA AGG GCT AAA CTG ATT TCC TTG CCA ACT AGT AAA AAC TGG ACC      2355
261 Thr Glu Arg Ala Lys Leu Ile Ser Leu Pro Thr Ser Lys Asn Trp Thr
262          765          770          775
264 TTC GGA CCC CAG GAC GTC GAT GAA CTG ATC TTC ATG CTG AGC GAC AGC      2403
265 Phe Gly Pro Gln Asp Val Asp Glu Leu Ile Phe Met Leu Ser Asp Ser
266          780          785          790
268 CCT GGG GTC ATG TGC CGG CCT TCC CGA GTC AAG CAG ATG TTT GCC TCC      2451
269 Pro Gly Val Met Cys Arg Pro Ser Arg Val Lys Gln Met Phe Ala Ser
270          795          800          805
272 AGA GCC TGC CGG AAG TCG GTG ATG ATT GGG ACT GCT CTT AAC ACA AGC      2499
273 Arg Ala Cys Arg Lys Ser Val Met Ile Gly Thr Ala Leu Asn Thr Ser
274 810          815          820          825
276 GAG ATG AAG AAA CTG ATC ACC CAC ATG GGG GAG ATG GAC CAC CCC TGG      2547
277 Glu Met Lys Lys Leu Ile Thr His Met Gly Glu Met Asp His Pro Trp
278          830          835          840
280 AAC TGT CCC CAT GGA AGG CCA ACC ATG AGA CAC ATC GCC AAC CTG GGT      2595
281 Asn Cys Pro His Gly Arg Pro Thr Met Arg His Ile Ala Asn Leu Gly
282          845          850          855
284 GTC ATT TCT CAG AAC TGACCGTAGT CACTGTATGG AATAATTGGT TTTATCGCAG A      2651
285 Val Ile Ser Gln Asn
286          860
288 TTTTATGTT TTGAAAGACA GAGTCTTCAC TAACCTTTTT TGTTTTAAAA TGAAACCTGC      2711
289 TACTTAAAAA AAATACACAT CACACCCATT TAAAAGTGAT CTTGAGAACC TTTTCAAACC      2771
W--> 290      2771
292 (2) INFORMATION FOR SEQ ID NO: 2:
294 (i) SEQUENCE CHARACTERISTICS:
295 (A) LENGTH: 862 amino acids
296 (B) TYPE: amino acid
297 (C) STRANDEDNESS: single
298 (D) TOPOLOGY: linear
300 (ii) MOLECULE TYPE: protein
301 (v) FRAGMENT TYPE: internal
303 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
305 Met Glu Arg Ala Glu Ser Ser Ser Thr Glu Pro Ala Lys Ala Ile Lys
306 1          5          10          15
307 Pro Ile Asp Arg Lys Ser Val His Gln Ile Cys Ser Gly Gln Val Val
308 20          25          30
309 Leu Ser Leu Ser Thr Ala Val Lys Glu Leu Val Glu Asn Ser Leu Asp
310 35          40          45
311 Ala Gly Ala Thr Asn Ile Asp Leu Lys Leu Lys Asp Tyr Gly Val Asp
312 50          55          60
313 Leu Ile Glu Val Ser Asp Asn Gly Cys Gly Val Glu Glu Glu Asn Phe
314 65          70          75          80
315 Glu Gly Leu Thr Leu Lys His His Thr Ser Lys Ile Gln Glu Phe Ala
316 85          90          95
317 Asp Leu Thr Gln Val Glu Thr Phe Gly Phe Arg Gly Glu Ala Leu Ser
318 100         105         110
319 Ser Leu Cys Ala Leu Ser Asp Val Thr Ile Ser Thr Cys His Ala Ser

```

VERIFICATION SUMMARY

DATE: 10/03/2005

PATENT APPLICATION: US/09/558,149

TIME: 10:58:38

Input Set : N:\Crf3\RULE60\09558149.raw

Output Set: N:\CRF4\09292005\I558149.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:36 M:238 W: Alpha Fields not Ordered, Reordered [(C) CLASSIFICATION:] of (1)(vi)
L:290 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1